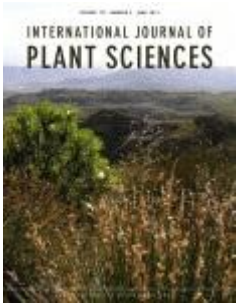

Morphological Variation in the Wild-Weedy Complex of Sorghum bicolor In Situ in Western Kenya: Preliminary Evidence of Crop-to-Wild Gene Flow?

James A. Okeno, Evans Mutegi, Santie de Villiers, Jeffrey D. Wolt and Manjit K. Misra

Vol. 173, No. 5 (June 2012), pp. 507-515

DOI: http://dx.doi.org/10.1086/665266

MORPHOLOGICAL VARIABILITY IN THE WILD-WEEDY COMPLEX OF *SORGHUM BICOLOR IN SITU*: PRELIMINARY EVIDENCE OF CROP-TO-WILD GENE FLOW IN WESTERN KENYA

J. A. Okeno, [1,*] E. Mutegi, [2,‡] S. de Villiers, [§] J. D. Wolt, [*] and M. K. Misra [*]

---

[*]Seed Science Center/Biosafety Institute for Genetically Modified Agricultural Products (BIGMAP), Iowa State University, Ames, IA 50011; [†]Kenya Agricultural Research Institute (KARI), National Gene Bank of Kenya, P.O. Box 30148, Nairobi, Kenya; and [§]International Crops Research Institute for the Semi-Arid Tropics (ICRISAT-Nairobi), P.O. Box 39063-00623, Nairobi, Kenya.

*[1] Author for correspondence; e-mail: jaketch@iastate.edu

[2] Current address: Department of Evolution, Ecology, and Organismal Biology, Ohio State University, 318W. 12th Ave., Columbus, OH 43210-1293

Abstract

Gene flow between cultivated plants and their wild/weedy relatives plays an important role in structuring the genetic variability within and among populations. The consequences of gene flow can contribute to the scientific basis (risk assessment) for managing agricultural systems, understanding evolutionary processes and designing *in situ* conservation measures for genetic resources and using these resources to secure current and future plant breeding programs. We conducted surveys and collections of wild and weedy sorghums in Lambwe valley in western Kenya, to investigate the genetic diversity within the wild-weedy complex of *Sorghum bicolor in situ*. We also attempted to understand the role, if any, of farmer practices and agro-climatic conditions on gene flow and genetic diversity. The morphological data presented here showed wide variability within wild-weedy sorghum populations with respect to habitats. "True wild" sorghum populations in national parks and a sugarcane belt were clearly distinguishable from the "putative hybrids" or intermediate forms found in sorghum fields, sorghum field margins and to some extent by the roadsides near sorghum fields. The existence of these intermediate forms is empirical evidence of introgression between cultivated sorghum and its wild-weedy relatives.

Keywords: Gene flow, morphological diversity, farmers' practices, introgression, *Sorghum bicolor*.

Introduction

Sorghum bicolor (L.) Moench is a highly diverse species, which belongs to the genus *Sorghum* of the tribe *Andropogoneae.* It has been divided into three subspecies, namely, *bicolor, verticilliflorum* and *drummondii* (Dogget 1988). Subspecies *bicolor* is recognized as consisting of five basic race, *bicolor, kafir, caudatum, durra* and *guinea*, and an additional ten intermediate races (Harlan and de Wet 1972). The four wild races or congeners *verticilliflorum*, *arundinaceum*, *virgatum* and *aethiopicum* are recognized in the subspecies *verticilliflorum* (Dogget 1988).Ssubspecies *drummondii* is a heterogeneous group composed of all the intermediate forms between wild and cultivated sorghums across the African continent. Sorghum is believed to have originated in northeast Africa, where it was domesticated about 3000 to 5000 years ago (Mann et al. 1983; Ejeta and Grenier 2005). It has been proposed that modern sorghums have diverse origins, with the cultivated subspecies *bicolor* arising from the wild subspecies *verticilliflorum.* It is thought that wild race *aethiopicum* gave rise to durra and bicolor cultivated races, while wild races *arundinaceum* and *verticilliflorum* gave rise to guinea and kafir types of sorghum, respectively (Mann et al., 1983). De Wet and Huckabay (1967); and De Wet et al. (1970), however, were of the opinion that durra sorghum arose from kafir.

Sorghum's closest wild relatives or congeners *S. bicolor* ssp. *verticilliflorum* and *S. bicolor* ssp. *drummondii*, introduced above are also native to Africa and are sexually compatible and naturally hybridize with its cultivated form (Doggett and Prasada Rao 1995). Sorghum is also interfertile with *S. propinquum* of southeastern Asia and *S. halepense*, which is native to southwestern Asia and adjacent North Africa (ref?). Both of these species are perennials, capable of vegetative propagation, dispersal and persistence with well-developed rhizomes (Ellstrand

2003). Both *S. propinquum* and *S. halepense* have also spread to other continents as weeds, where S. *halepense*, johnsongrass, is described as one of the world's worst weeds (Holm et al. 1977). Progeny segregation, allozyme and molecular analyses have revealed presence of crop-specific alleles in populations of wild *S. bicolor* when it co-occurs with the crop in Africa, suggesting intraspecific hybridization and introgression (Doggett and Majisu 1968; Aldrich and Doebley 1992; Aldrich et al. 1992). Molecular analysis has also confirmed hybridization between cultivated sorghum (*S. bicolor*) and johnsongrass (*S. halepense*) in the United States (Paterson et al. 1995; Arriola and Ellstrand 1997; Morrell et al. 2005). Spontaneous hybridization between cultivated sorghum and these wild relatives has been reported at varying rates (Arriola and Ellstrand 1996; Pedersen et al. 1998). Arriola and Ellstrand (1996) reported spontaneous hybridization between johnsongrass (*Sorghum halepense*) and cultivated sorghum as far as 100 m apart with frequencies as high as 2%. Pedersen et al. (1998) reported an average out-crossing rate of 48% between cultivated sorghum and sudangrass (*Sorghum bicolor* ssp *drummondii*). In Africa, pollen flow from cultivated to wild-weedy sorghums was predicted to occur naturally at frequencies of 2.5% at a distance of 13m (Schmidt and Bothma 2006). Recent surveys in Ethiopia and Niger (Tesso et al. 2008) and in Kenya (Mutegi et al. 2009) showed that sorghum congeners are found intermixed with and adjacent to cultivated sorghum and that their flowering periods overlapped with that of the cultivated sorghum. Mutegi et al. (2009) further showed morphological evidence of hybridization within the *Sorghum bicolor* species at a country scale in Kenya.

Sorghum is a vitally important crop in Africa and much of the developing world. It has a remarkable ability to endure both drought conditions and water-logging and it grows well on

marginal lands. It is the dietary staple of more than 500 million people in more than 30 countries with only rice, wheat, maize, and potatoes feeding more people than sorghum. Due to this importance, there is a resurgence in efforts to create genetically engineered (GE) sorghum, with improved agronomic and quality traits following the successful recovery of transgenic plants either through *Agrobacterium*-mediated or particle bombardment systems. Zhu et al. (1998); Krishnaveni et al. (2000) transferred the agronomically important rice chitinase gene, *chi II*, under the constitutive CaMV 35S promoter, into sorghum in an attempt to breed for resistance to stalk rot (*Fusarium thapsinum*). Tadesse et al. (2003) generated transgenic sorghum plants expressing the *dhdps-raec1* mutated gene encoding an insensitive form of the dihydrodipicolinate synthase, the key regulatory enzyme of the lysine pathway, in an attempt to generate sorghum lines with elevated lysine content. Zhao et al. (2003) transformed sorghum with an *Agrobacterium*-mediated transformation system with a binary vector containing a DNA insert containing the HT12, alpha hordothionin, a high lysine-coding gene from barley. Gao et al. (2005) transformed sorghum with an agronomically useful rice *tlp*-D34 gene for resistance to fungal pathogens and drought. Transgenic sorghum plants have also been produced for resistance against neonate larvae of the spotted stem borer *(Chilo partellus)* via particle bombardment of shoot apices expressing a synthetic *cry1Ac* gene from *Bacillus thuringiensis* (*Bt*) under the control of a wound-inducible promoter from the maize protease inhibitor gene (*mpiC1*) (Girijashankar et al. 2005). Sorghum lines sampled from Kenya were transformed with *HarChit* and *HarCho* antifungal genes from the fungus *Trichoderma harzianum* using particle bombardment (Ayoo et al 2008). Lately, sorghum has been engineered under the Africa Biofortified Sorghum (ABS) Project (www.supersorghum.org), to improve grain digestibility,

increase grain content of essential amino acids lysine, threonine and tryptophan, and improve bioavailability of iron and zinc (Zhao et al. 2008).

Since some of the above efforts may lead to the first commercial release into the environment of GE sorghum, regions in Africa and other developing world, countries with high diversity of crop sorghum landraces and their wild relatives are of special interest and raises considerable ecological concerns to regulators, environmentalists and policy makers due to potential pollen-mediated gene flow from GE sorghum to its wild relatives. Tanksley and McCouch (1997) considered such regions as "genetic insurance" where alleles lost through domestication and modern breeding can be recovered by falling back on the crop wild relatives or ancestors. Thus, gene flow plays a role in structuring the genetic variability within and among populations and understanding its consequences can contribute to scientifically based risk assessment for managing agricultural systems, understanding evolutionary processes and designing *in situ* conservation measures for genetic resources and using these resources to secure current and future plant breeding programs as described in other studies (Schmidt and Bothma 2006; Barnaud et al. 2008; Mutegi et al. 2009; Hokanson et al, 2010). Mutegi et al. (2009) suggested that special efforts should be directed to record and map wild sorghum populations in Kenyan national parks, as a possible further evidence to estimate the extent and direction of historical and recent gene flow between cultivated and wild sorghum for contribution to the national genetic resource conservation policy.

We therefore chose to conduct surveys and collections of wild and weedy sorghums within Ruma National Park and adjacent farms in Lambwe valley and also in a sugarcane belt in Migori

District of Suba District, both of greater South Nyanza, western Kenya. This area is known to host many wild-weedy sorghum taxa. In this region, farmers mainly carry out traditional farming methods and maintain their own sorghum landraces. Since it was important not only to study gene flow based on biological traits but also to gather knowledge on farmer practices, we also collected samples of cultivated sorghum varieties and recorded farmer knowledge of these varieties. This information is important to allow for the investigation of the role, if any, of agro-climatic and farmer practices on gene flow and genetic diversity in the crop-wild-weedy complex of *Sorghum bicolor in situ*. Data on the movement of genes from cultivated sorghum to its co-occurring wild-weedy relatives in cultivated and natural habitats is important in the assessment of potential ecological risk for introducing GE sorghum near or around the Ruma National Park. Currently, there is no information available on where such studies have been conducted. The objectives of this study were to 1) carry out local scale surveys and collections of wild and cultivated sorghum samples from Ruma National Park and adjacent farms in Suba District and in a sugarcane belt in Migori District of greater South Nyanza, western Kenya, that host crop-wild-weedy sorghum complexes, and 2) estimate the extent and direction of historical and recent gene flow between cultivated and wild sorghum at the agro-ecosystem and/or local scale. This paper reports on the analysis of morphological data while the molecular data is being developed and will be presented in another manuscript.

Material and methods

Study/sampling area

Field surveys were carried out in Lambwe valley of Suba District that houses Ruma National Park and its adjacent farms. Ruma National Park was gazetted in 1966, as Lambwe Valley Game

reserve, and acquired National park status in 1983. It lies on the flat floor of Lambwe Valley, surrounded by settled lands with a mix of small-scale cultivation of various crops, including sorghum, and grassy pastureland. The park is a valuable island of natural habitat in a sea of human settlement. Like other protected lands it is an important reservoir for plant genetic resources, especially crop-wild relatives (CWR). To provide a contrast with a sorghum farming system, the survey was extended into the Awendo sugarcane belt in the neighboring Migori District.

Our study aimed to acquire samples of wild sorghum from the park for genetic analysis in comparison with wild-weedy sorghum samples obtained from different habitats within and around sorghum fields. An exploratory survey was first conducted in the target region during the last week of June 2009. The purpose of the survey was to (i) ascertain the presence of wild sorghum within Ruma National Park, (ii) identify the optimum time for simultaneously collecting samples of cultivated and wild sorghum since the latter tends to mature earlier and (iii) identify and map potential sites for collecting populations of cultivated sorghum and its wild-weedy relatives. Subsequently, a sample collection trip was conducted in the first two weeks of July 2009.

Various habitats were selected for crop-wild gene flow analysis and comparison among populations of cultivated and wild sorghum with different levels of spatial overlap (table 1). The habitats were classified and abbreviated as: (i) sorghum fields (sf) where wild types were co-occurring with cultivated counterparts, (ii) sorghum field margins (fm), in close proximity (less than 5m) to cultivated sorghum, (iii) disturbed ground by the roadside (rd) but close to cultivated

sorghum (5-10m), (iv) sugarcane fallow (sc) in a zone where sorghum is not grown, and (v) protected land in a national park/wildlife sanctuary (pk). The SF represented habitats of complete crop-wild intermix, whereas fm and rd represented intermediate habitats close to farmlands. The pk represented a natural habitat isolated from cultivation, while sc represented farmland habitat away from sorghum cultivation. In total 8 populations of wild-weed sorghum were sampled in the above-mentioned 5 contrasting habitats from collection sites, S1, S2, S3, S4, S5, S7, and S8 (see table 1). All collection sites within each habitat were geo-referenced using a hand-held GPS and location coordinates used to generate a collection map using DIVA GIS (fig. 1).

Table 1

Geographical coordinates of the study wild-weedy sorghum populations and sampling sites

No.	Population code	Sampling site	Latitude (°S)	Longitude (°W)
1	SF1	S1	-0.55642	34.29691
2	SF2	S2	-0.55796	34.28625
3	RD	S2	-0.55796	34.28625
4	FM	S3	-0.66900	34.05918
5	PK1	S4	-0.62407	34.26198
6	PK2	S5	-0.58695	34.26075
7	SF3	S7	-0.72887	34.07066
8	SC	S8	-0.93209	34.55222

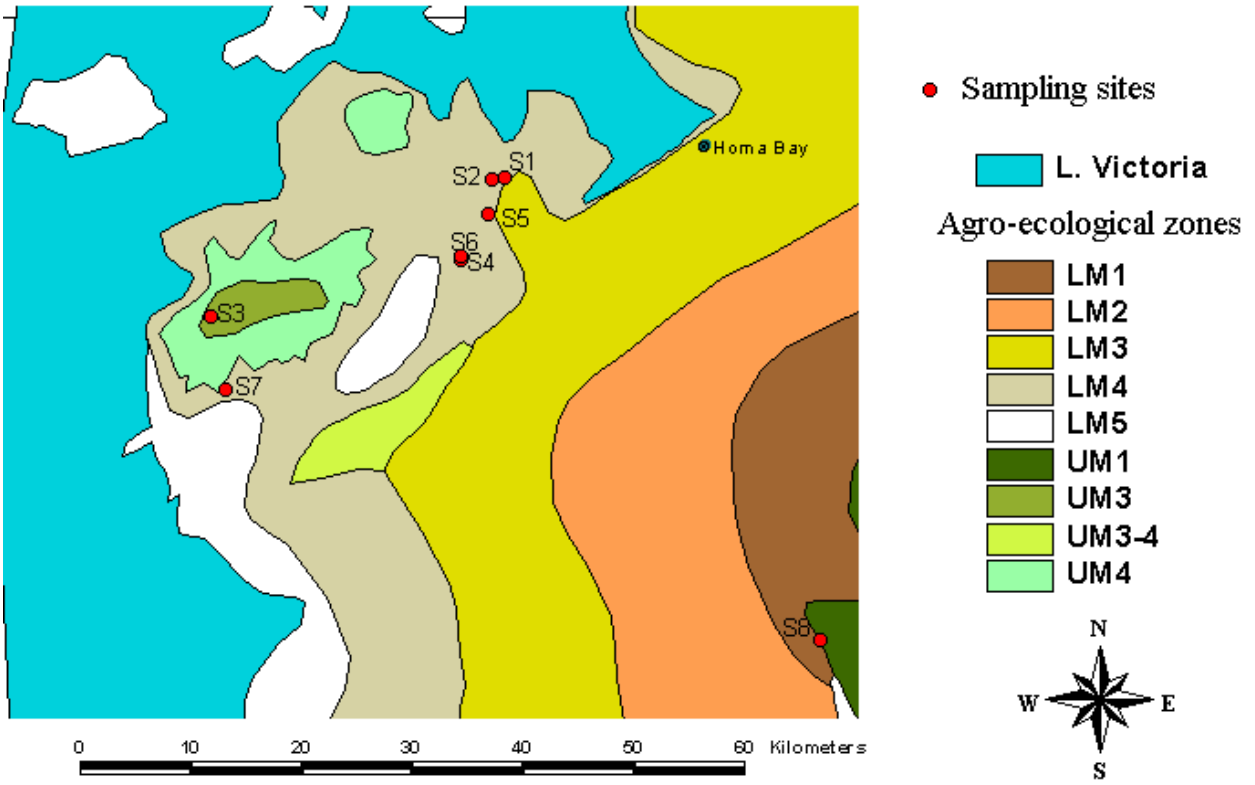

Fig. 1 Map of the study area with sampled sites (S1 – S8) and agro-ecological zones shown. The abbreviation LM and UM denote lower midland and upper midland, respectively.

In addition to the wild populations, a total of 13 populations of cultivated sorghum, representing 6 distinct farmer-named landraces from 5 different farmers' fields were also sampled (table 2). All except two of the cultivar populations (*nyar koyoko* and *kumba*) had at least one corresponding sympatric wild sorghum population. For each population of cultivated sorghum (farmer variety) and its wild relative, seeds were sampled from 25 randomly selected individual maternal plants, located at least 1m apart.

To investigate morphological variability within and between the wild-weedy sorghum populations, five random plants were selected in each sampled population and data recorded on five quantitative (plant height, number of internodes, number of basal tillers, panicle length,

panicle width) and five qualitative (panicle shape, presence/absence of awn, glume color, glume cover) traits. Each of the five individuals in each population was categorized as either "true wild" (w) or "putative hybrid" (h) using morphological observation as described previously by Dogget (1988).

Table 2

Traditional names, botanical race, seed color, and sampling site of sorghum landraces collected

No	Traditional name	Race	Seed color	Sampling site	Farmer field
1	*andiwo rachar*	Caudatum	White	S1	1
2	*andiwo rabuor*	Caudatum	Brown	S1	1
3	*ochuti*	Durra/Durra-Caudatum	Red	S2	2
4	*serena*	Kaffir-Caudatum	Red	S2	2
5	*andiwo rachar*	Caudatum	White	S2	2
6	*andiwo rabuor*	Guinea-Caudatum	Brown	S2	2
7	*andiwo rabuor*	Guinea-Caudatum	Brown	S3	3
8	*kumba*	Guinea-Caudatum	White	S3	3
9	*andiwo rabuor*	Caudatum	Brown	S3	3
10	*nyar koyoko*	Durra-Caudatum	Brown	S6	4
11	*kumba*	Guinea-Caudatum	White	S6	4
12	*oboke nyar tende*	Guinea-Caudatum	Dark brown	S7	5
13	*ochuti*	Durra-Caudatum	Brown	S7	5

Examples of the two types of wild-weedy sorghum based on panicle characteristics are shown in figure 2. “Putative hybrids” were characterized by compact to semi-compact panicles (fig. 2a; 2b), while “true wilds” had open and loose panicle types (fig. 2c; 2d).

Fig. 2 Wild-weedy sorghum panicles. Type a and b are typical of “putative hybrids” and were mostly found in sorghum fields, while type c and d are typical of “true wild” with the former being found in the park and the latter in a sugarcane field**.**

Statistical analysis

All data was analysed using the software R.2.11.1 (R Development Core Team 2008). Basic descriptive statistics (mean, range and standard deviation) were calculated for the five quantitative traits based on the entire data set. Uni-variate analysis of variance (ANOVA) was performed on the five quantitative traits to explore the level of variation among the eight

populations. Principal component analysis (PCA) was performed on a data matrix of the 5 quantitative traits and 40 representative individuals in order to explore patterns of association and major traits contributing to the delineation. Input data was standardized to minimize the effect of scale on variability weighting. Patterns of grouping among individuals were visualized by plotting the first two principal components of the PCA. A hierarchical cluster analysis was further performed based on Gower's dissimilarity coefficient (Gower 1971) as implemented in the algorithm DAISY (Kaufman and Rousseeuw 1990) of the package CLUSTER in R. This procedure enabled generation of a pairwise dissimilarity matrix for cluster analysis using both qualitative and quantitative data. The dissimilarity matrix was further subjected to analysis of similarity (ANOSIM), implemented in the package VEGAN of R to compare patterns of variation within and among populations, within and among habitats and within and between wild-weedy types (h:"putative hybrid"; w: "true wild") for multiple morphological characters. ANOSIM tests for differences between groups in a manner analogous to ANOVA by comparing within-group (population or habitat) similarity to between-group similarity, with p-values determined randomly (Legendre and Legendre 1998). The distances in the dissimilarity matrix are converted to ranks, so that the smallest distance ranks (r) as 1. The ANOSIM statistic R is thus based on differences of mean ranks between groups (rB) and within groups (rW):

$$R = (rB - rW)/(N(N-1)/4)$$

where N is the total number of individuals. Values of R range from 0 (no difference between groups; i.e. null hypothesis) to 1 (complete divergence between samples). Significance testing was obtained using 10,000 permutations.

Results

Wide variability is found within the studied wild-weedy sorghum populations for all quantitative traits measured (table 3). Variation was also highly significant (P≤0.001) among populations for these traits. Patterns of morphological variability are further summarized and compared among the eight wild-weedy populations using box plots (fig. 3a – fig. 4c).

Table 3

Mean, minimum, maximum, standard deviation (SD) and F-test significance (outcome of ANOVA) for the five quantitative traits recorded for 40 individuals of wild-weedy sorghum in 8 populations

Trait	Min	Max	Mean	SD	[1]F-test
Plant height (cm)	189.0	381.0	282.1	59.73	***
No. of internodes	5.00	15	8.7	2.81	***
Panicle length (cm)	8.00	49	31.1	9.59	***
Panicle width (cm)	6	40	21.1	10.51	***
No. of tillers	0	15	3.4	3.48	***

[1] ANOVA among 8 wild-weedy populations

*** Highly significant (P≤0.001)

Generally, populations of wild-weedy sorghum found in cultivated sorghum fields (sf) were taller in height than their counterparts in the non-sorghum growing habitats (fig. 3a). A similar trend was observed for number of inter-nodes (fig. 3b). Furthermore, populations of wild-weedy sorghum co-occurring with cultivated sorghum were characterized by panicles that are shorter (fig. 4a) and more compact (fig. 4b) than their counterparts in the park or sugarcane field. The images presented in figure 2 support the boxplot results for panicle characteristics. Populations

of wild-weedy sorghum occurring in field margins (fm) and roadside (rd) habitats appeared to be morphologically intermediate between those in sorghum fields (sf) and those in non-cultivated sorghum habitats (pk, sc). No consistent patterns were observed among the populations for the number of basal tillers (fig. 4c).

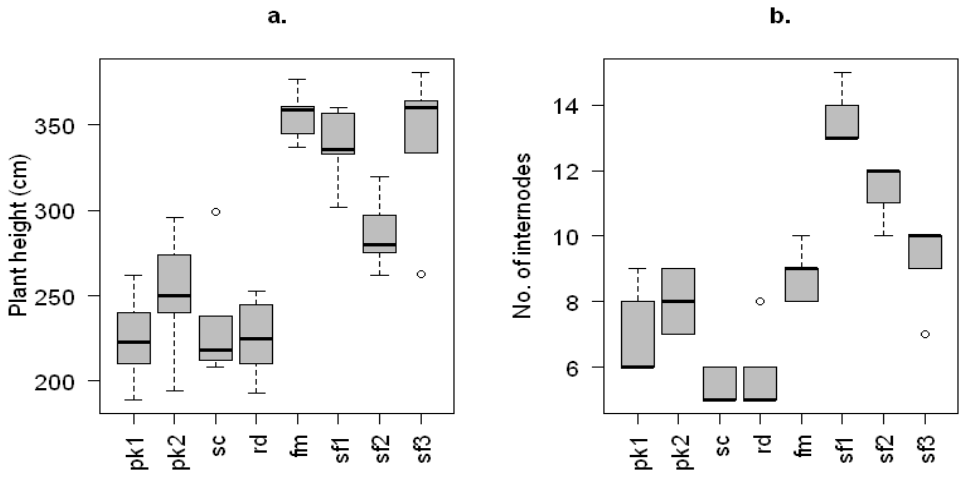

Fig. 3 Box plot showing variability in height (a) and number of internodes (b) for the seven populations of wild-weedy sorghum in this study. The box presents the inter-quartile range (50% 0f values); while the line across the box indicates the median. The lines running vertically from the box (whiskers) extend to the highest and lowest values, excluding outliers which are denoted by circles.

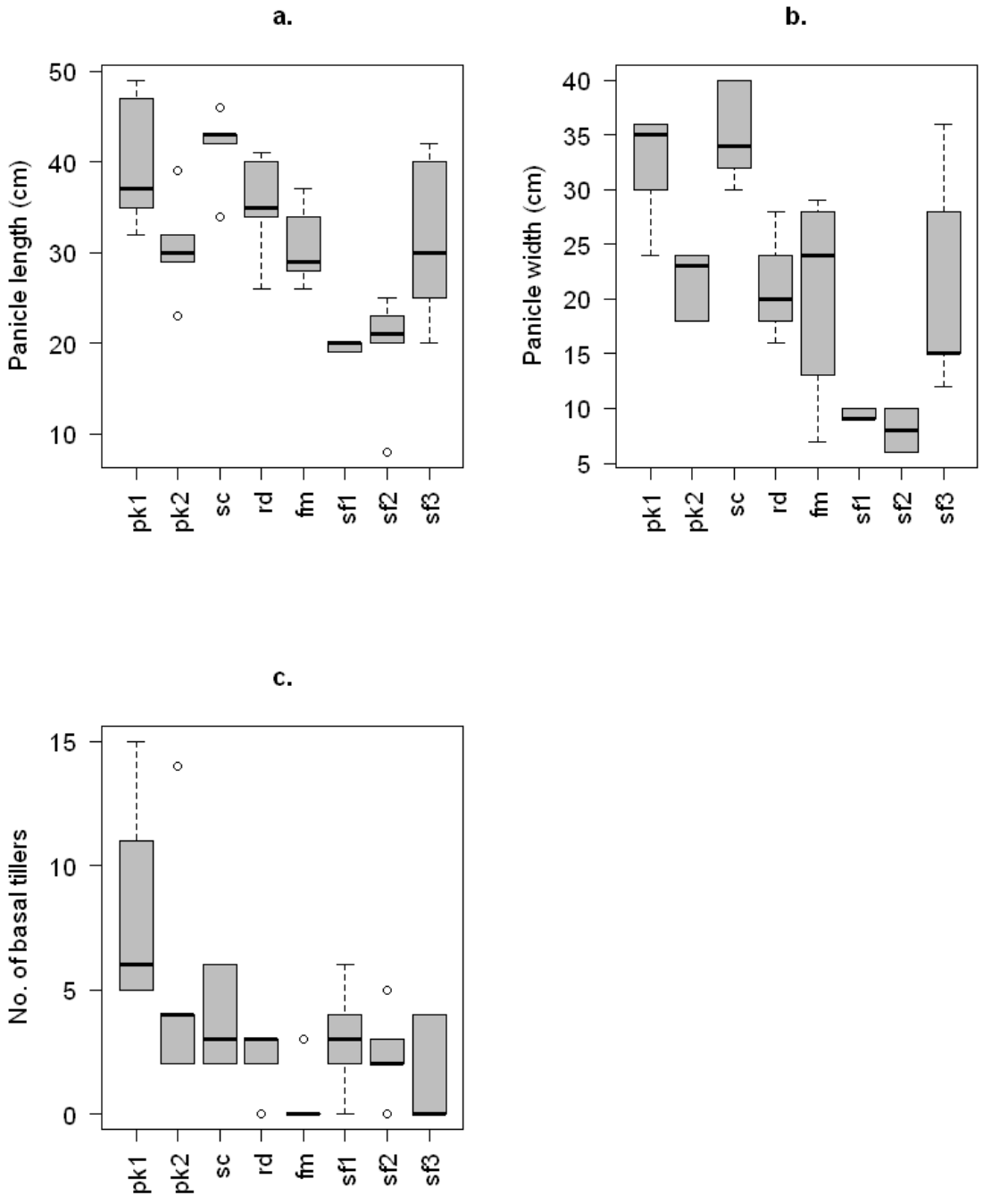

Fig. 4 Box plot showing variability in panicle length (a) and panicle width (b) and number of basal tillers for the seven populations of wild-weedy sorghum in this study. The box presents the inter-quartile range (50% of values); while the line across the box indicates the median. The lines running vertically from the box (whiskers) extend to the highest and lowest values, excluding outliers which are denoted by circles

Specific patterns that defined the way the five quantitative traits were associated to influence the components of the PCA are presented in table 4. The first three principal components (PCs) explained over 90% of the variation among individuals. The first component, PC1, explained more than half (56%) of the variability, with high negative loadings from plant height and number of internodes, and high positive loadings from panicle length and panicle width. The second principle component explained 19.1% of the total variability and was mainly described by the number of basal tillers with high negative loadings. The third principle component accounted for 15.8% of the total variability and was most heavily weighted by plant height, panicle length and panicle width.

Table 4

Principle components, eigenvalues, component loadings, and amount of total variance explained in a principal components analysis on five quantitative morphological traits

	PC1	PC2	PC3	PC4	PC5
Eigen-value	1.65	0.96	0.88	0.54	0.39
Component loadings					
Plant height	-0.42	0.15	0.76		-0.48
No. of internodes	-0.53	-0.27	0.24		0.77
No. of basal tillers	0.23	-0.93	0.17	-0.13	-0.21
Panicle length	0.49	0.22	0.43	-0.65	0.32
Panicle width	0.50		0.39	0.75	0.20
Propotion of variance	0.56	0.19	0.16	0.06	0.03
Cumulative proportion of variance	0.56	0.75	0.91	0.97	1.00

A bi-plot of PC1 and PC2 further revealed two major morphotypes. Individuals identified in the field as putative hybrids were clearly separated from those classified as “true” wilds by PC1 with negative and positive scores, respectively (fig. 5). The putative hybrids were characterized by tall plants with high number of internodes and with panicles that were largely short and compact. Contrastingly, plants categorized as “true” wilds were generally shorter, with low number of

internodes and with panicles that were generally long and broad. Furthermore, putative hybrids appeared to separate into two groups, one made up exclusively of individuals from sorghum fields and another made up of all individuals from field margins and a few of those from sorghum fields.

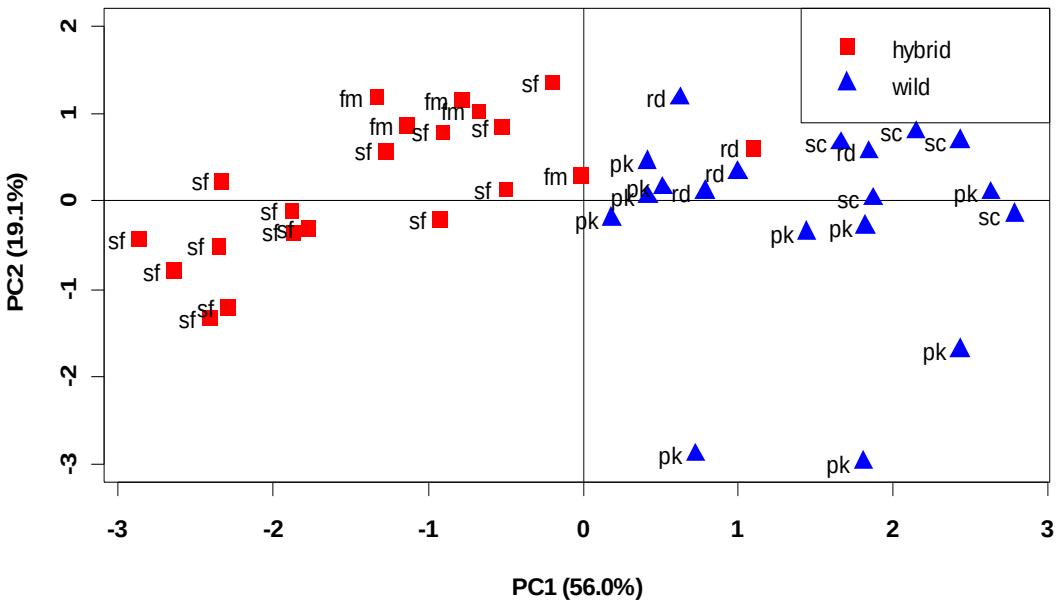

Fig. 5 Bi-plot of the first two principal components based on 5 quantitative traits recorded. PC1 and PC2 explain 56.0% and 19.1% of the total variability among individuals, respectively. Source habitats are abbreviated as follows: sorghum field (sf), sorghum field margin (fm), roadside (rd), sugarcane field (sc) and national park (pk).

Cluster analysis

Hierarchical cluster analysis revealed a high level of variability among the wild-weedy sorghum individuals, with an overall dissimilarity level slightly below 80% (fig. 6). Three major clusters were evident at approximately 56% level of dissimilarity, with a clear delineation between "putative hybrids" and "true wilds". Cluster A was exclusive to putative hybrid individuals from sorghum fields, whereas cluster B contained mostly putative hybrids from intermediate habitats

(fm and rd). All except two members of Cluster C were of the "true wild" category and were drawn largely from habitats well isolated from sorghum fields (pk and sc), even though about 20% proportion of individuals in this group were from intermediate habitats (fm and rd). Overall, there was close congruence between PCA and cluster analysis results. ANOSIM revealed highly significant morphological variability among populations ($R=0.76$; $P<0.001$), habitats ($R=0.48$; $P<0.001$) and wild-weedy sorghum type ($R=0.57$; $P<0.001$).

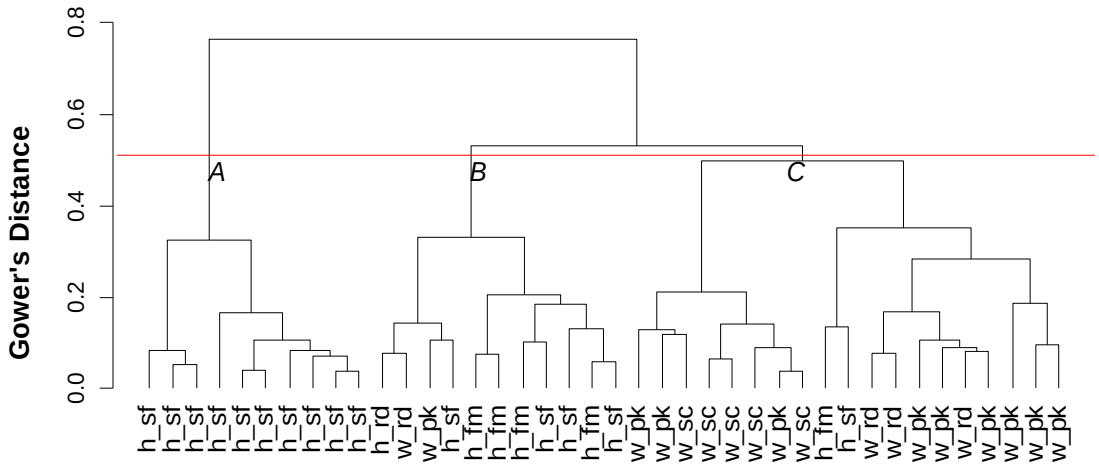

Fig. 6 Dendogram of the hierarchical cluster analysis performed on 40 wild and weedy sorghum and 5 quantitative traits based on a pairwise Gower's dissimilarity coefficient as labeled by habitat (1, 2, and 3 representing three morphological groupings described under "Discussion")

Discussion

The morphological traits of the wild and weedy sorghums were observed to vary according to habitat and also with proximity to cultivated sorghum. The morphological appearance of wild-weedy individuals was used to categorize the populations as either "putative crop-wild hybrid" or "true wild". The PCA analysis revealed that plant height, number of internodes and panicle

length and width as the most important morphological traits distinguishing between “putative hybrids” and “true wild”. Both hierarchical cluster and PCA analyses clearly grouped together individual plants found within or around the sorghum fields (“putative hybrids”) and those from national park, roadsides, and the sugarcane belt (“true wild”). Putative crop-wild hybrids were generally vigorous with characters observed to be intermediate between cultivated and wild sorghum i.e. large grains, reduced or totally missing awn, less clasping glumes, glume color similar to some of the cultivated sorghum varieties, broad leaves, semi-compact to compact panicles, high level of shattering, prolific seed producers and few or no basal tillers. Within the national park the wild sorghum populations occupied mostly disturbed areas by roadsides and below trees, whereas in the sugarcane belt they occupied sugarcane fallows. At least six attributes classified plants as “true wilds” i.e. shorter plant height, narrow leaves, prominent awns, panicles that are open and often with very loose and drooping primary branches, glumes that cover the grain fully with a tight clasp, and large number of basal and secondary tillers. These findings indicate that conservation efforts for wild sorghum should target natural habits like national parks and other areas far away from sorghum cultivation fields. Dogget (1988) provided a clear distinction between the cultivated and wild sorghum forms in East Africa in his remarks that, “On sites such as abandoned cultivation, evident hybrids may be seen, sometimes also in farmers' fields. These usually have more closed panicles than the wild type, with broader leaves, larger grain and tight black glumes. The grains shed readily when ripe, with the glume attached. These hybrids cannot be confused either with the cultivated range of material, or with the wild type”. To further strengthen our findings, wild sorghums with characteristically small stature, thin culm, and very loose panicles were also found in abandoned fields, un-weeded fields, field crop margins and by the roadsides in earlier investigations (Teso et al. 2008; Mutegi

et al. 2009). Although we cannot determine the extent and direction of gene exchange between cultivated and wild sorghum at this juncture based on our morphology data, the existence of intermediate forms ("putative hybrids") in sorghum fields, sorghum field margins and to some extent by the roadside near sorghum fields is empirical evidence that introgression has been occurring between cultivated sorghum and its wild-weedy relatives in this ecosystem. Similar conclusions were made based on morphological data at national (Ejeta and Grenier 2005; Teso et al. 2008; Mutegi et al. 2009) and local or village level (Barnaud et al. 2009).

The dynamics of gene introgression between cultivated and wild/weedy sorghum can be highly influenced by farmer practices (Barnaud et al. 2009). The farmers in Lambwe valley, Suba District, belong to the *Luo* tribe, sub-tribe *Abasuba,* who practice traditional farming methods. During the crop growth cycle, hand-weeding is carried out either once or twice, depending on the size of the field and labor available per household. In our investigation, we found weedy and cultivated sorghums sympatric with each other, with weedy sorghums either in sorghum fields or by field margins. The weedy sorghums were of two morphotypes, *ogolo* or *magolo* (in sorghum fields) and *oboro* (by field margins or roadsides), identified and named so by the *Abasuba*. The farmers select against these types and remove them from sorghum fields if they are identified during weeding. However, farmers concurred that it can be a daunting task to distinguish these weedy sorghums from cultivated sorghums before flowering.

Farmers easily distinguish weedy sorghums which persist to harvesting time and leave them standing in the fields. After harvest, the farmers would graze their cattle on crop residues including weedy types. In the process, the weedy sorghums would be further dispersed to other

crop fields, open fallows, field margins or roadsides. Thus, cattle in this area play an important role in dissemination and survival of putative hybrids like *ogolo* (*magolo)* or *oboro* outside farmers' field. We also observed that farmers in this ecosystem maintain their own sorghum landraces, originally received from a relative of the same village or from another village, while a few were being purchased from local markets. These different sorghum landraces (see table 2) were often found grown mixed in the same farmer's field. Molecular data would be able to show us how different sorghum landraces from the same field contributed their genes to the crop-wild/weedy genepool, thereby also contributing to the dynamics of introgression.

Acknowledgments

Funding from the Biosafety Institute for Genetically Modified Agricultural Products (BIGMAP), Iowa State University enabled participation of J.A. Okeno in this study. Also appreciated is Dr. Kendall Lamkey, chair, Department of Agronomy, Iowa State University for a travel grant to J.A. Okeno for a reconnaissance trip to in the Lambwe Valley, Kenya. We are also grateful to Zachary Muthamia for logistical support from the National Gene Bank of Kenya. We acknowledge permits and security accorded by the Kenya Wildlife Services (KWS) to enter and collect wild sorghum in Ruma National Park. Last, but not least, we indebted to Jonathan Hirsi, Charles Ndiege and Richard Otip for participating in all sample collection trips.

Literature cited

Aldrich PR J Doebley 1992 Restriction fragment variation in the nuclear and chloroplast genomes of cultivated and wild *Sorghum bicolor*. Theor Appl Genet 85:293–302.

Aldrich PR, J Doebley, KF Schertz, A Stec 1992 Patterns of allozyme variation in cultivated and wild *Sorghum bicolor*. Theor Appl Genet 85:451–460.

Arriola PE, N Ellstrand 1996 Crop-to-weed gene flow in the genus *Sorghum* (Poaceae): Spontaneous interspecific hybridization between johnsongrass, *Sorghum halepense* and crop sorghum, Sorghum bicolor. Am J Bot 83:1153–1160.

Arriola PE, N Ellstrand 1997 Fitness of interspecific hybrids in the genus sorghum: persistence of crop genes in the wild populations. Ecol Appl 7:512–518.

Ayoo L, M Bader, D Becker, H Lorz 2008 Genetic Transformation of Kenyan Sorghum (Sorghum bicolor L. Moench) with Chitinase and Chitosanase for Fungal Disease Resistance. PhD Thesis University of Hamburg, Germany.

Barnaud A, G Triguersos, D McKey, HI Joly 2008 High outcrossing rates in fields with mixed sorghum landraces: how are landraces maintained? Heredity 101:445–452.

Barnaud A, M Deu, E Garine, J Chantereau, J Bolteu, EO Koida, D McKey, HI Joly 2008 A weed – crop complex in sorghum: the dynamics of genetic diversity in a traditional farming system. Am J Bot 96:1869–1879.

De Wet JMJ, JP Huckabay 1967 Origin *of Sorghum bicoIor*. II. Distribution and domestication. Evolution: 21:787– 802.

De Wet JMJ, JR Harlan, EG Price 1970 Origin of variability in the spontanea complex of sorghum bicolor. Am J Bot 57:704–707.

Doggett H 1988 Sorghum, 2nd edn. John Wiley and Sons, New York.

Doggett H, BN Majisu 1968 Disruptive selection in crop development. Heredity 23:1–26.

Doggett H, KE Prasada Rao 1995 Sorghum. Pages 173–180 *in* J Smartt, NW Simmonds, eds. Evolution of crop plants. Harlow: Longman. 2nd edn.

Ejeta G, C Grenier 2005 Sorghum and its weedy hybrids. Pages 123–135 *in* J Gressel ed. Crop ferality and volunteerism. Boca Raton, FL: CRC Press.

Ellstrand N 2003 Dangerous Liaisons? When cultivated plants mate with their wild relatives. Johns Hopkins University Press, Baltimore & London.

Gao Z, X Xie, Y Ling, S Muthukrishnan, GH Liang 2005 Agrobacterium tumefaciens-mediated sorghum transformation using a mannose selection system. Plant Biotechnol J 3:591–599.

Girijashankar V, HC Sharma, KK Sharma, V Swathisree, LS Prasad, BV Bhat, M Royer, BS Secundo, ML Narasu, I Altosaar, N Seetharama 2005 Development of transgenic sorghum for insect resistance against the spotted stem borer (*Chilo partellus*) Plant Cell Rep 24:513–522.

Gower JC 1971 A general coefficient of similarity and some of its properties. Biometrics 27:623–637.

Harlan JR, JMJ de Wet 1972 A simplified classification of cultivated sorghum. Crop Sci 12:172–1756.

Holm LG, DL Plucknett, JV Poncho, JP Herberger 1977 The world's worst weeds: distribution and biology. Uni. Press of Hawaii, Honolulu.

Kaufman L, PJ Rousseeuw 1990 Finding Groups in Data: An Introduction to Cluster Analysis. Wiley, New York.

Krishnaveni S, JM Jeoung, S Muthukrishnan, GH Liang 2000 Transgenic sorghum plants constitutively expressing a rice chitinase gene show improved resistance to stalk rot. J Genet Breed 55:151–158.

Legendre P, L Legendre 1998 Numerical Ecology. 2nd English Edition, Elsevier, Amsterdam.

Mann JA, CT Kimber, FR Miller 1983 The origin and early cultivation of sorghums in Africa. Bulletin no. 1454. Texas A and M University, College Station, Texas, USA.

Morrell PL, TD Williams-Coplin, AL Lattu, JE Bowers, JM Chandler, AH Paterson 2005 Crop-to-weed introgression has impacted allelic composition of johnsongrass populations with and without recent exposure to cultivated sorghum. Mol Ecol 14:2143–2154.

Mutegi E, F Sagnard, M Muraya, B Kanyenji, B Rono, C Mwongera, C Marangu, J Kamau, H Parzies, S de Villiers, K Semagn, PS Traoré, M Labuschagne 2009. Ecogeographical distribution of wild, weedy and cultivated Sorghum bicolor in Kenya: Implications for conservation and crop-to-wild gene flow. Genet Resour Crop Evol DOI 10.1007/s10722-009-9477-7.

Paterson AH, KF Schertz, YR Lin, SC Liu, YL Chang 1995 The weediness of wild plants: Molecular analysis of genes influencing dispersal and persistence of johnsongrass, Sorghum halepense (L.) Pers Proc Nat Acad Sci USA 92:6127–6131.

Pedersen JF, JJ Troy, B Johnson 1998 Natural outcrossing of Sorghum and Sudangrass in the Central Great Plains. Crop Sci 38:937–939.

R Development Core Team 2008 R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Schmidt M, G Bothma 2006 Risk assessment for transgenic sorghum in Africa crop-to-crop gene flow in Sorghum bicolor (L.) Moench. Crop Sci 46:790–798.

Tadesse Y, L Sági, R Swennen, M Jacobs 2003 Optimisation of transformation conditions and production of transgenic sorghum (*Sorghum bicolor*) via microparticle bombardment. Plant Cell Tissue Organ Culture 75:1–18.

Tanksley, SD , SR McCouch 1997 Seed banks and molecular maps: Unlocking genetic potential from the wild. Science 277:1063–1066.

Tesso T, I Kapran, C Grenier, A Snow, P Sweeney, J Pedersen, D Marx, G Bothma, G Ejeta 2008 The potential for crop-to-wild gene flow in sorghum in Ethiopia and Niger: A geographic survey. Crop Sci 48:1425–1431.

Zhao ZY, K Glassman, V Sewalt, N Wang, M Miller, S Chang, T Thompson, WE Catron, D Bidney, Y Kebede, R Jung 2003 Nutritionally improved transgenic sorghum. Pages 413–416 *in* Plant Biotechnology 2002 and Beyond.

Zhao ZY, R Jung, K Glassman, R Chikwamba, L Mehlo, N Mkhonza, M O'Kennedy, A Grootboom, G Beyene, C Erasmus, J Taylor, J Mutisya, K Mburu, P Anderson 2008 The Africa Biofortified Sorghum – Applying biotechnology to develop nutritionally improved sorghum for Africa. Page 102 in Abstract 1st All Africa Congress on Biotechnology, 22-26 September 2008, Nairobi, Kenya.

Zhu H, S Muthukrishnan, S Krishnaveni, G Wilde, JM Jeoung, GH Liang 1998 Biolistic transformation of sorghum using a rice chitinase gene. J Genet Breed 52:243–252.